Several kinds of tRNA genes of *Xenopus laevis* are clustered together within tandemly repeated 3.18-kilobase DNA fragments. Other members of these reiterated tRNA gene families are dispersed and irregularly arranged in the genome. Here we report the isolation and some characteristics of one such dispersed gene that codes for a tyrosine tRNA. It is located within a low copy number 9.4-kilobase restriction fragment that contains no other RNA polymerase III gene functional *in vitro*. The dispersed gene differs from the clustered tyrosine tRNA gene by a single purine transition within the coding region, by extensive sequence differences within the intervening sequence and 5' and 3' flanking regions, and by its ~6-fold higher transcriptional activity in homologous S-100 extracts. Analyses of hybrid genes and deletion mutants demonstrate that this differential transcription is due to DNA in the 5' flanking regions.

The tRNA genes of the frog *Xenopus laevis* are unusual in two respects; they are more highly reiterated than in any other examined eukaryote, and they exhibit a unique form of organization. Once again, sea urchins provide a precedent in several interesting respects. The question of how such complex expression patterns are regulated is of considerable interest.

Both DNA strands in no obviously ordered way. Of these eight genes, only that encoding tRNA<sup>Tyr</sup> contains an intervening sequence.

Although most members of these reiterated tRNA gene families appear to reside within this major cluster, a few copies are found in other locations. For example, a minor fraction of the initiator methionyl-tRNA hybridization is found outside the major 3.18-kb cluster when genomic *X. laevis* DNA is fractionated on actinomycin-CsCl gradients (Clarkson et al., 1978). In more informative assays, the entire 3.18-kb repeat or subclones containing some of its tRNA genes have been used as hybridization probes against restriction enzyme digests of *X. laevis* genomic DNA. A minor class of tandem repeats has been detected in this way. They possess the same length as the major 3.18-kb repeat but only a subset of its genes, and they are present in only 12–15 copies/haploid genome (Rosenthal and Doering, 1983). The same kind of assay has also revealed several other restriction fragments, ranging in size from 1.4 to over 23 kb and present in even fewer copies, that hybridize to single DNA gene probes (Rosenthal and Doering, 1983; Postel et al., 1984).

These dispersed tRNA genes may represent portions of the major repeat cluster that have recently been translocated to other regions of the genome. Examples of such “orphans” have been found in the early histone gene family of sea urchins. Most members of this family are located in tandem repeats that contain all five histone genes, but the occasional single gene, together with some of its spacer DNA, has also been found embedded in nonhistone sequences at remote genomic locations (Childs et al., 1981). The in vitro function, if any, of these early histone gene orphans is unknown.

A more interesting possibility for the dispersed *X. laevis* tRNA genes is that their expression is regulated during development and that this is somehow aided by their nonclustered organization. Once again, sea urchins provide a precedent, for their histone genes active late in embryogenesis are present in low copy number and are widely dispersed rather than clustered (Maxson et al., 1983).

For these reasons we have been interested in examining a *X. laevis* tRNA gene that resides outside the major 3.18-kb cluster. Here we report the isolation and characterization of a dispersed tyrosine tRNA gene that differs from its clustered counterpart in several interesting respects. The question of its *in vivo* function remains unresolved, but it is transcribed much more actively *in vitro* than the clustered tRNA<sup>Tyr</sup> gene, a property that reflects sequence differences in the DNA upstream of the two genes.

**EXPERIMENTAL PROCEDURES**

**DNA Isolation—**Genomic *X. laevis* DNA was prepared from nucleated blood cells of a single frog as described by Birnstiel et al. (1972). Standard methods (Maniatis et al., 1982) were used to prepare DNA from single plaques of the *λ* vector Charon 30 (Rimm et al., 1978).
junctions of all hybrid and deletion mutant clones were determined vectors but with 378-bp filled-in MspI-ApaI inserts. Sequences at the large EcoRI-PvuII fragment of pTyrC*; the reciprocal combination HindIII site and BamHI site of pBR327. The hybrid gene plasmid was filled in and the fragment then inserted between the filled-in EcoRI linkers were added, and the fragment was ligated between the ApaI and filled-in EcoRI sites of pTyrD. The comparable filled-in BanI-ApaI fragment from pTyrC* inserted into the same two sites of pBR322. The tRNAPhe gene was subcloned separately on PstI site of pBR322. The tRNAPhe gene and its 3' flanking region (Fig. 4). Other single tRNA replaced by the comparable MspI-BglII fragment from the dispersed DNA for in vitro transcription and restriction mapping. Subcloning and DNA Sequence Analysis—The following subclones were prepared from λ21B, a phage recombinant containing a single copy of the major 3.18-kb repeat unit (Clarkson et al., 1978). The tRNATyr gene and tRNAC* genes present on a 422-bp HapI fragment (Müller and Clarkson, 1980) were inserted via dC-dG tailing into the PstI site of pBR322. The tRNAψ gene was subcloned separately on a 386-bp HindIII-BglI fragment; the BglI site was made flush ended. EcoRI linkers were added, and the fragment was ligated between the EcoRI and HindIII sites of pBR327. This subclone is denoted pPhE (Fig. 1). A subclone of the tRNAψ gene was similarly prepared by ligating a 264-bp Hpal fragment between filled-in HindIII and BamHI sites of pBR327. This subclone is denoted pTyrC (Fig. 1 and pTyrC* thereof to distinguish this clustered gene from the dispersed tRNAψ gene. Subclone pTyrC* is identical except that the DNA encoding the intragenic MspI site was replaced by the comparable MspI-BglII fragment from the dispersed tRNAψ gene and its 3' flanking region (Fig. 4). Other single tRNA gene subclones from the 3.18-kb repeat have been described previously; pMet-A contains an initiator methionine tRNA gene (Koski and Clarkson, 1982). Phage were screened by the method of Benton and Davis (1977) with a nick-translated subclone containing the tRNAψ and tRNAψ* genes from the 3.18-kbp repeat. Positive clones were plaque purified and were used to prepare DNA for in vitro transcription and restriction mapping. Detergent Analyses—S-100 extracts were incubated with several restriction enzymes that lack sites within the 3.18-kb repeat. The resulting fragments were fractionated on 1% agarose gels, denatured, transferred to nitrocellulose filters (Southern, 1978), and hybridized with a nick-translated (Maniatis et al., 1975) subclone containing the tRNAψ and tRNAψ* genes from the repeat. Our results are in excellent agreement with those of Rosenthal and Doering (1983). With all enzymes tested, the majority of the hybridization is found in high molecular weight DNA corresponding to the intact major 3.18-kb repeat cluster. A significant amount of hybridization also occurs with the minor 3.18-kb repeats that possess single sites for BamHI, PstI, and XhoI. In addition, BamHI generates at least 10 weakly hybridizing bands that range from <2 to >15 kb; their intensities, in comparison with that of the minor 3.18-kb repeats, suggest that each is present in 1–3 copies/haploid genome (Rosenthal and Doering, 1983; data not shown).

Molecular Cloning of Dispersed tRNA Genes—To try to clone these weakly hybridizing bands, genomic DNA was first partially digested with BamHI. DNA fragments of 6-15 kb were then selected by sucrose gradient centrifugation and were packaged in vitro into the λ phage vector Charon 30 (Rimm et al., 1980). To achieve the desired recombination between any repeated sequences that may be present within the inserts, the phage were propagated in an E. coli 803 recA strain. This has the disadvantage that viable phage need one copy of the "stuffer fragment" carrying the red and gan genes, thereby reducing the capacity of this vector. A total of 8.7 × 104 phage were probed with the nick-translated subclone containing the tRNAψ and tRNAψ* genes, resulting in 13 positives. Two of these clones, λX30 and λX74, exhibited particularly strong hybridization and were selected for further study. Restriction analysis showed that each phage clone contains one copy of the Charon 30 "stuffer fragment" and a single BamHI fragment of X. laevis DNA. To improve DNA yields, the 9.6-kb insert from λX30 was recloned in the BamHI site of pBR327 to give the plasmid clone pX30. The 9.4-kb insert of λX74 was similarly recloned, yielding pX74.

In Vitro Transcription of Plasmid Clones—To determine if either clone contains a functional tRNA gene, the supercoiled plasmid DNAs were separately incubated with S-100 extract of X. laevis tissue culture cells that contain RNA polymerase III and the necessary co-factors for specific transcription

3 A. Mazalbreaud, D. Scherle, D. Runnger, and S. G. Clarkson, manuscript in preparation.
Structure and Transcription of Dispersed X. laevis tRNA\textsuperscript{\textsc{tyr}} Gene

(Weil et al., 1979). Clone pXI30 is completely inert in these extracts (Fig. 1), suggesting that it contains a nonfunctional pseudogene or homologies with spacer DNA rather than with gene sequences. This clone has not been investigated further. Clone pXI74, on the other hand, is transcribed more actively than any of the single tRNA gene subclones from the 3.18-kb repeat (Fig. 1). After a 1-h reaction at 20 °C, its major products comprise two intense but slightly heterogeneous RNA bands of ~100 and ~96 nucleotides. During the same incubation time there is significant conversion of the tRNA\textsuperscript{\textsc{tyr}} and tRNA\textsuperscript{\textsc{lys}} precursors to mature length tRNA (Koski and Clarkson, 1983; Fig. 1). However, the in vitro tRNA\textsuperscript{\textsc{tyr}} gene transcripts are not efficiently processed, thereby suggesting that the S-100 extracts are defective in splicing activity under these incubation conditions. The paucity of pXI74 transcripts in the mature tRNA region of the gel provided the first indication that this clone also contains a tRNA gene with an intervening sequence.

Transcription Mapping—To locate the tRNA gene(s) within the 9.4-kb insert of pXI74, the sites for several restriction enzymes were mapped by single and double digestions and by partial digestion of end-labeled fragments (Smith and Birnstiel, 1976). The pXI74 DNA was then digested with several of these enzymes, and the resulting fragments were gel fractionated, denatured, transferred to a nitrocellulose filter, hybridized with the DNA, and autoradiographed. The fragments that hybridize are indicated by arrows on the photograph of the gel. M, a double digest of a DNA with BglII and HindIII, with fragment lengths given in kb.

sequences are thus delimited to a region of ~500 bp of DNA (Fig. 3).

DNA Sequencing—The nucleotide sequence of this region was then determined by the chemical method (Maxam and Gilbert, 1980). As indicated by the sequencing strategy (Fig. 3), each strand of the 560 bp of DNA between the sites for HindIII and BglII was sequenced in its entirety. The sequence (Fig. 4) reveals, between nucleotides 96 and 180, a putative tRNA\textsuperscript{\textsc{tyr}} gene with the following interesting features. 1) Its 5' and 3' flanking regions are very different from those surrounding the tRNA\textsuperscript{\textsc{tyr}} gene on the 3.18-kb repeat. 2) It contains a 12-bp rather than 13-bp intervening sequence, of which only 4 bp are identical in the two genes. 3) It also differs from the clustered tRNA\textsuperscript{\textsc{tyr}} gene by a single G→A transition at position 57 in the TΨ loop. This single base change was independently confirmed by restriction mapping because it results in the loss of a HindIII site.

In Vitro Transcription of Plasmid Subclones—The 560 bp of sequenced DNA between the HindIII and BglII sites (Fig. 3) was subcloned into pBR327. This subclone, denoted pTyrD, generates the same pattern of in vitro transcripts as the parent clone containing a 9.4-kb insert (compare Figs. 1 and 5A). This strongly suggests that the putative tRNA\textsuperscript{\textsc{tyr}} gene is responsible for the observed transcription in vitro and that it is the only functional tRNA gene within the original cloned fragment of 9.4 kb.

This dispersed tRNA\textsuperscript{\textsc{tyr}} gene has the further interesting property that it seems to be transcribed much better than the clustered tRNA\textsuperscript{\textsc{tyr}} gene (Fig. 1). Precise quantitation is difficult because the presumed termination site of the clustered gene, TTTGTTTTT in the noncoding strand (Muller and Clarkson, 1980; Fig. 4), appears to be inefficient in vitro,
Fig. 3. Restriction and transcription map of the pX174 insert and the subsequent sequencing strategy. The location of various restriction sites within the 9.4-kb insert is shown in the upper part of the figure. Thick lines denote fragments that hybridize to the major pX174 in vitro transcript (Fig. 2 and data not shown). The sequencing strategy is expanded in the lower part of the figure. 3'-Labeled (solid lines) or 5'-labeled (dashed lines) fragments were sequenced by the chemical method from the indicated restriction sites.

Fig. 4. Comparison of the DNA sequences within and around the dispersed and clustered tRNA\textsuperscript{\textsc{Ty}} gene\textsuperscript{s}. The noncoding strand is shown of the dispersed gene, TyrD (this work) and of the clustered gene, TyrC (Müller and Clarkson, 1980). Nucleotides encoding the mature tRNAs are presented in bold type. The location of the intervening sequences is shown by a dashed line and the intragenic base change by an asterisk. The restriction sites at the sequence boundaries were used to construct subclones pTyrD and pTyrC. For pTyrC*, the 3' flanking region of pTyrC was replaced with that of pTyrD by sequence exchange at the indicated MspI site. The hybrid genes 5'-C-3'D and 5'-D-3'C were constructed by similar exchanges at the indicated PvuII site.

resulting in read-through transcripts of variable length (Fig. 1). To test this proposition and, in so doing, to try to circumvent the problem, the 3' flanking region of the clustered gene was replaced with that of the dispersed gene. The latter contains a stretch of 15 contiguous T residues (Fig. 4) that is expected to constitute a highly efficient termination site (Bogenhagen and Brown, 1981). The replacement was made at the MspI site located just downstream of the intragenic base change (Fig. 4). The resulting subclone, pTyrC*, thus contains the 5' flanking region and the entire coding region of the clustered tRNA\textsuperscript{\textsc{Ty}} gene joined to the 3' flanking region of the dispersed gene. The expected junction was confirmed by DNA sequencing (Fig. 5E).

The very long in vitro transcripts characteristic of the unmanipulated clustered gene are not made by subclone pTyrC*. Instead, it produces a major RNA band that is slightly heterogeneous in length and both shorter and much less intense than the pTyrD transcripts (Fig. 5A). These results imply that the T\textsubscript{tr} tract is indeed an efficient terminator for both pTyrD and pTyrC* transcripts and that the T-rich region normally located behind the clustered gene is only a moderately effective terminator in vitro. More important, they strongly suggest that the transcripts are derived from the tRNA\textsuperscript{\textsc{Ty}} gene within each subclone and that these two genes functionally differ either in their transcriptional activities or in the stabilities of their transcripts.

**RNA Sequencing of in Vitro Transcripts**—To define the transcriptional units of both genes, the major in vitro transcripts of pTyrD and pTyrC* (indicated by arrows in Fig. 5A) were recovered from preparative reactions containing \([\alpha-\text{\textsuperscript{32P}}]\text{GTP}, [\alpha-\text{\textsuperscript{32P}}]\text{UTP}, or [\alpha-\text{\textsuperscript{32P}}]\text{ATP}. After RNase T1 digestion, the resulting oligonucleotides were fingerprinted (Fig. 5, B-D) and were then identified by secondary analyses with RNase A, RNase T2, and, in some cases, nuclease P1 (Table I). As summarized in Fig. 5E, the sequence and molar yields of the oligonucleotides match those predicted from the DNA sequences within and immediately around the two genes.

The dispersed tRNA\textsuperscript{\textsc{Ty}} gene starts transcription at two sites. The major site, accounting for \(-80\%\) of the initiation events, is the G residue located 7 nucleotides in front of the mature tRNA-coding sequence. The remaining transcripts (indicated by small arrows in Fig. 5A) start with a pppAp at position -3. In contrast, the clustered gene initiates with a unique purine, in this case a pppAp at position -5. Termination of all three kinds of transcripts occurs at the second, third, and fourth T residue within the T\textsubscript{tr} stretch behind each gene. The dispersed gene thus gives rise to 100–102 and 96–98-nucleotide-long tRNA precursors that differ by 4 nucleotides at their 5' ends and by a variable number of 3'-terminal U residues. The pTyrC* transcripts comprise a related set of 99–101-nucleotide-long precursors that exhibit the same heterogeneity at their 3' ends.
In Vitro Transcription of Hybrid Genes—The preferential accumulation of TyrD precursors (Fig. 5A) implies either that the dispersed gene is transcribed more efficiently or that these transcripts are more stable than those made by the clustered gene. A difference in transcriptional activity between the two genes could be due to the single intragenic base transition or the different intervening sequences or flanking regions (or any combination thereof). Altered transcript stability could be a reflection of any of these sequence differences when they are present in the two kinds of precursors. To try to distinguish between these various possibilities, the two genes were cleaved at the PvuII site located within the T9 loop does not inhibit transcription nor does it lead to instability of the transcripts. The presence of the 13-bp intervening sequence from the clustered gene is also without effect on transcription activity or transcript stability. As expected, however, it does lengthen the in vitro 5′-3′ hybrid transcripts by one nucleotide. Similarly, the 5′-3′ hybrid appears to be transcribed no better than the clustered gene, thereby confirming that high transcriptional activity is not dependent on either the intragenic base change or the intervening sequence of the dispersed gene. More important, the dispersed gene transcripts accumulate preferentially over those of the 5′-3′ hybrid (Fig. 6). The only difference between these two genes is the nature of their 5′ flanking sequences. These sequences could act at the DNA level to cause the two genes to be transcribed with different efficiencies. Alternatively, the different 5′ leader sequences could alter the stabilities of the two kinds of tRNA precursors in the cell-free extracts.

To distinguish between these two possibilities, the primary transcripts of the dispersed gene and of the 5′-3′ hybrid were isolated by gel electrophoresis and were incubated for various times with fresh S-100 extract. The two kinds of precursors are equally stable under standard transcription reaction conditions (data not shown). After a 1-h incubation, ~10% of the transcripts are found as shorter fragments of variable length that may represent processing intermediates. However, most of the primary transcripts of both genes are still intact. The preferential accumulation of the dispersed gene transcripts is thus the result of an event at the DNA rather than RNA level. Specifically, 5′ flanking DNA sequences interact with the transcription machinery to bring about the differential expression of the dispersed and clustered tRNA<sup>Tyr</sup> genes.

In Vitro Transcription of 5′ Flanking Sequence Deletion Mutants—To try to delimit these upstream sequences, deletions were introduced into the 5′ flanking regions of both pTyrD and pTyrC*. Because such manipulations are more properly considered as deletion substitutions, each deletion break point was ligated to two different heterologous DNA segments. The junctions were made at either the ClaI or EcoRI sites of the pBR327 vector in such a way that the plasmid sequences preceding these sites replaced the natural 5′ flanking DNA. In view of the recent finding that high template DNA concentrations obscure the effects of some deletion

**Fig. 5.** In vitro transcripts of pTyrD and pTyrC* and the sequences of their RNase T1 products. A, the indicated plasmid DNAs were incubated for 1 h at 20 °C in X. laevis S-100 extracts with [γ-<sup>32</sup>P]GTP; RNA was extracted, fractionated on a 12% polyacrylamide-8.3 M urea gel, and autoradiographed. The X. laevis tRNA<sup>21</sup> subclone (Fig. 1) was included in each reaction as an internal standard; the lengths of its in vitro transcripts are given in nucleotides. RNAs isolated for sequence analysis are indicated by arrows. B, RNase T1 fingerprint of the major pTyrD transcript labeled with GTP. Its specific oligonucleotides are numbered. The 5′ end (pppGpG, spot D4) was electrophoresed too far in this particular fingerprint. C, RNase T1 fingerprint of GTP-labeled pTyrC* transcripts. Its specific oligonucleotides are numbered. D, composite diagram of all fingerprints. Oligonucleotides specific to the dispersed gene are indicated with diagonal stripes, while those specific to the clustered gene are stippled. E, the RNase T1 products of the major transcripts are aligned with their RNA and noncoding DNA strand sequences. Common oligonucleotides are shown only below the TyrD RNA sequence. The numbering corresponds to that in Table I. DNA encoding the mature tRNAs is emphasized by bold type, and the intragenic base change by an asterisk above the TyrD DNA sequence.
TABLE I
Sequence of RNase T1 products of the in vitro transcripts

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common products:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Gp (U,A,C)</td>
<td>-</td>
<td>Gp</td>
<td>Gp</td>
</tr>
<tr>
<td>2</td>
<td>CGp (G)</td>
<td>Cp</td>
<td>Gp</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>CUGp (G)</td>
<td>Up, Gp</td>
<td>Cp</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>CUCGp (U)</td>
<td>Cp</td>
<td>Cp</td>
<td>Gp</td>
</tr>
<tr>
<td>9</td>
<td>ACUGp (U)</td>
<td>Up</td>
<td>ApCp, Gp</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>CUCGp (C)</td>
<td>ApGp</td>
<td>Cp</td>
<td>Cp</td>
</tr>
<tr>
<td>13a</td>
<td>ACUOH</td>
<td>-</td>
<td>ApCp, Up</td>
<td>-</td>
</tr>
<tr>
<td>13b</td>
<td>ACUUOH</td>
<td>-</td>
<td>ApCp, Up</td>
<td>-</td>
</tr>
<tr>
<td>13c</td>
<td>ACUUUOH</td>
<td>-</td>
<td>ApCp, Up</td>
<td>-</td>
</tr>
<tr>
<td><strong>Products specific to dispersed gene:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2a</td>
<td>UUCAAUCCGp (G)</td>
<td>Cp, Gp</td>
<td>ApAUp, Up</td>
<td>ApAUp, Cp</td>
</tr>
<tr>
<td>D2b</td>
<td>TUCAm1AUCCGp (G)</td>
<td>Cp, Gp</td>
<td>Apm1AUp, Tp</td>
<td>Apm1AUp, Cp</td>
</tr>
<tr>
<td>D4</td>
<td>pppGp (A)</td>
<td>pppGp</td>
<td>-</td>
<td>pppGp</td>
</tr>
<tr>
<td><strong>Products specific to clustered gene:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>UGp (U,A)</td>
<td>Up</td>
<td>Gp</td>
<td>Gp</td>
</tr>
<tr>
<td>C2</td>
<td>AUGp (A)</td>
<td>Cp</td>
<td>ApUp</td>
<td>Gp</td>
</tr>
<tr>
<td>C3</td>
<td>UUCGp (A)</td>
<td>Cp</td>
<td>Up</td>
<td>Gp</td>
</tr>
<tr>
<td>C4</td>
<td>CUCUGp (A)</td>
<td>Cp</td>
<td>Cp, Up</td>
<td>Gp</td>
</tr>
<tr>
<td>C5</td>
<td>AUCUGp (G)</td>
<td>Cp, Gp</td>
<td>ApUp</td>
<td>-</td>
</tr>
<tr>
<td>C6</td>
<td>pppACCpGp (G)</td>
<td>Cp</td>
<td>-</td>
<td>pppApCp</td>
</tr>
</tbody>
</table>

*a Numbered as in Fig. 5.

*b Deduced from the position of the oligonucleotides in the fingerprints, the mobilities, and molar yields of the RNase A and T2 products, and the DNA sequences. Nearest neighbor bases are indicated in parentheses.

*c Columns list the labeled RNase A products. Asterisks denote the labeled phosphate groups after RNase T2 digestion.

- not present.

5'-End groups determined by digestions with RNase T2 and nuclease P1.

substitutions (Wilson et al., 1985), the mutants were tested at 0.2, 2, and 10 μg/ml template DNA supplemented, where necessary, with pBR322 DNA to a final DNA concentration of 10 μg/ml. Identical results were obtained under all three conditions.

These deletion substitutions affect both the frequency and position of transcription initiation. The overall transcriptional activity of the dispersed gene, for example, is reduced by ~40% in the 5'-12a mutant (Fig. 7). This mutant makes use of the same two initiation sites as the wild-type gene, and in approximately the same ratio, but a minor fraction of the transcripts initiate at a new site. From RNase T2 analysis of transcripts labeled with [α-32P]UTP (data not shown), this site is identified as the A at position -16. The 5'-12a mutant of the clustered gene, in which transcription is reduced by ~50%, also makes use of this minor start site. As with the wild-type genes, in these two mutants the dispersed gene is transcribed ~6 times better than the clustered gene, thereby implying that the DNA lying just upstream of the mature coding sequence is the most important for determining the frequency of initiation events.

Comparison of the 5'-12b mutants suggests, however, that this is too simplistic a conclusion. These mutants utilize a minor new start site, a G at -24 or -25, and the dispersed gene is again preferentially transcribed. But the difference in transcriptional activity of the two genes is only 3-fold, and the A at -3 is now the major initiation site for transcripts of the dispersed gene (Fig. 7). Hence, sequences upstream of -12
492

Structure and Transcription of Dispersed X. laevis tRNA<sup>Tyr</sup> Gene

![Image](https://example.com/image.png)

**Fig. 6. In vitro transcription of hybrid genes.** The indicated plasmid DNAs were incubated for 1 h at 20°C in X. laevis S-100 extracts with [α-<sup>32</sup>P]GTP; RNA was extracted, fractionated on a 12% polyacrylamide-8.3 M urea gel, and autoradiographed. The X. laevis tRNA<sup>Tyr</sup>-<sup>es</sup> subclone (Fig. 1) was included in each reaction as an internal standard; the lengths of its in vitro transcripts are given in nucleotides.

We document here the existence of a X. laevis tyrosine tRNA gene that resides outside the major 3.18-kb repeat cluster. This gene is found on a 9.4-kb BamHI fragment which is not visible as a discrete band in genomic Southern blots that readily detect the 12–15 copies of a minor 3.18-kb repeat family (Rosenthal and Doering, 1983 and data not shown). We conclude that the 9.4-kb BamHI fragment is very rare, on the order of 1–3 copies/haploid genome. This fragment contains no other tRNA genes that are functional in vitro nor any other kind of gene transcribed by RNA polymerase III (Figs. 1, 2, 3, and 5). From its location within the 9.4-kb fragment, the tRNA<sup>Tyr</sup> gene is separated from its nearest possible tRNA gene neighbors by >5.7 kb on one side and by >3.5 kb on the other (Fig. 3). This dispersed tRNA<sup>Tyr</sup> gene is thus very different, in both copy number and gross organization, from its counterpart that is clustered together with seven other tRNA genes within the major 3.18-kb tandem repeat family.

The dispersed and clustered genes also differ at a detailed structural level (Fig. 4). They contain 12- and 13-bp intervening sequences, respectively, that exhibit little homology. The mature tRNA-coding regions, although highly conserved, are distinguished by a single purine transition. Finally, the two genes possess no significant homologies in their 5' or 3' flanking regions, apart from their T-rich termination sites. These sequence differences imply that the dispersed tRNA<sup>Tyr</sup> gene is not an example of an "orphon" (Childs et al., 1981), i.e. it does not simply represent a portion of a 3.18-kb repeat that has been moved to a new location. Rather, they suggest that this gene has been kept distinct from the repeats for a considerable length of time and that its coding sequence, at least, has been subjected to selective pressure.

To our knowledge, these two X. laevis tRNA<sup>Tyr</sup> genes represent the only characterized examples of vertebrate tRNA genes that are interrupted by intervening sequences. It is, therefore, all the more notable that the two introns do exhibit some sequence homology. The intervening sequence within the clustered gene, contrary to our earlier suggestion (Muller and Clarkson, 1980), is now known to be located one nucleotide 3' to the anticodon (Laski et al., 1982a; Filipowicz and Shatkin, 1983). This is the same position as in yeast tRNA genes (reviewed by Ogden et al., 1984), and we assume that it also applies to the intervening sequence within the dispersed gene. This location would mean that the two introns are...
identical for 4 nucleotides at their 3' ends.

We suspect that this short region of homology may be important for the efficiency of splicing in Xenopus because it would permit unspliced precursors from the two genes to adopt very similar secondary structures (compare A and C, Fig. 8). These structures, calculated from the rules of Tinoco et al. (1973), would allow the intervening sequences to base pair with only one nucleotide of the anticodon (or two nucleotides if G-U pairing is permitted). They are thus more "open" than the structures of yeast tRNA precursors, in which the intervening sequences are complementary with large portions of the anticodon loop and usually permit base pairing with all three nucleotides of the anticodon (Ogden et al., 1984). Hence, tRNA splicing endonucleases from various organisms may have slightly different preferences for certain structural features of their substrates. Consistent with this view, a mutation that weakens the complementarity between the anticodon of a yeast tyrosine tRNA and its intervening sequence results in more efficient splicing of the precursor in X. laevis oocytes (Nishikura et al., 1982). We note, however, that recent stability estimates (Cech et al., 1983) would permit the intervening sequence of the X. laevis dispersed tRNA\textsuperscript{\textgamma} gene, but not of the clustered gene, to fully base pair with the anticodon (Fig. 8).

A major functional difference between these two genes is the ~6-fold higher transcripational activity of the dispersed gene in homologous cell-free extracts (Figs. 1, 5, and 7). The hybrid gene constructs (Fig. 6) clearly demonstrate that this differential transcripational activity is not due to the single intragenic base change or to the different intervening sequences within the two genes. Neither conclusion is unexpected. First, the A-G transition at position 57 within the T9 loop, although comprising part of the essential intragenic "B block" promoter element (Galli et al., 1981), is found at a position invariably occupied by a purine. Second, \textit{in vivo} synthesis of a mature yeast suppressor tRNA\textsuperscript{\textgamma} is reduced by only 40% when the wild-type gene is replaced by a mutant lacking the intervening sequence; suppressor activity is inhibited more severely in this deletion mutant, but this seems to be due to the absence of a U→Ψ modification within the anticodon rather than to slightly diminished transcription (Johnson and Abelson, 1983).

The hybrid gene constructs demonstrate instead that the 5' flanking regions of the two genes are responsible for their differential expression \textit{in vitro} (Fig. 6). These results, together with those from simple deletion-substitution analysis (Fig. 7), are consistent with the large body of evidence indicating the importance of upstream sequences for both the position and frequency of eukaryotic tRNA gene transcription (see Wilson et al., 1985 for recent review). The way these sequences function, however, is far from clear. The dispersed tRNA\textsuperscript{\textgamma} gene continues to be transcribed better than the clustered gene in constructs that retain only 12 bp of normal 5' flanking DNA (Fig. 7), implying that these sequences at least contribute to the differential activity of the two genes. The 12 bp of DNA preceding the dispersed gene are relatively A+T rich (58%) whereas the comparable region ahead of the clustered gene is both very G+C rich (83%) and dyad symmetric. Of these two features, base composition seems to be the more important because the region of dyad symmetry is lengthened in one mutant but shortened in another, yet both mutants are transcribed with very similar efficiencies (compare pTyrC* 5'-12a and 5'-12b, Fig. 7). The enhanced activity of the clustered gene in mutant 5'-3a is also consistent with the proposal that transcription is inhibited when the initiation site is part of a very G+C rich region. However, this region is quite A+T rich in mutant 5'-3b; yet it has an activity even lower than that of the wild-type gene. Presumably this reflects the contribution of sequences even farther upstream. In keeping with the recent suggestion of Shaw and Olson (1984), unwinding of the DNA duplex in this mutant may be hindered by the G+C rich DNA between positions -21 and -25 (Fig. 7). It should be noted, however, that a collection of over 30 flanking sequence deletion mutants of a X. laevis tRNA\textsuperscript{\textmet} gene has revealed no simple correlation between transcriptional activity and base composition at any position upstream of the initiation sites (Hipikind and Clarkson, 1983). We thus conclude that the differential expression of the dispersed and clustered genes is in part due to sequences immediately around the anticodon rather than to slightly diminished transcription (Johnson and Abelson, 1983).

Another interesting and unresolved question is that of the \textit{in vivo} function of these two tyrosine tRNA genes. It is possible that one or the other represents a variant that is not normally expressed in Xenopus. This seems improbable in the case of the clustered gene because, after suitable mutagenesis

---

**Fig. 8. Secondary structure models of the dispersed and clustered tRNA\textsuperscript{\textgamma} gene transcripts.** A, a transcript of the X. laevis dispersed gene including its intervening sequence and mature 5' and 3' ends. The asterisk denotes the position of the G→A transition. B, an alternative secondary structure for the anticodon stem and loop and intervening sequence of the dispersed gene. C, secondary structure model of the comparable region of the X. laevis clustered tRNA\textsuperscript{\textgamma} gene. The anticodon is enclosed in brackets. Predicted splice sites are indicated by arrows. All nucleotides are shown in unmodified form.
of the anticodon, it is able to produce suppressor tRNAs that are functional in mammalian cells (Laski et al., 1982b, 1984). We also think it improbable in the case of the dispersed gene. Its high transcriptional activity in vitro is striking, as is the sequence conservation within the coding region and at the 3' end of the intervening sequence. Moreover, the mere presence of an intragenic base change does not necessarily imply non-functionality. For example, yeast contains two tRNA\textsuperscript{Tyr} genes that differ by an A-U transversion at position 59 of the T\textsuperscript{Ψ} loop yet they are equally active in vivo (Ogden et al., 1984). For all these reasons we think it likely that the dispersed tRNA\textsuperscript{Tyr} gene is expressed in Xenopus. One intriguing possibility is that it may represent an example of a somatic tRNA gene whereas the clustered gene, because of its high copy number and tandem organization, may be expressed during oogenesis. Experiments to test this idea are underway.

Acknowledgments—We are indebted to Janine Corlet for expert technical assistance.

REFERENCES